

A compilação de um
Corpus de Referência Literário
Lourdes Bernardes Gonçalves
UFC / USP



Grupo Temático: Lingüística de Corpus:
Um novo olhar para o ensino de LE e Tradução

Universidade Federal de Uberlândia
XI Simpósio Nacional de Letras e Lingüística
I Simpósio Internacional de Letras e Lingüística

22 a 24 de novembro de 2006



Definição e Metodologia

■ Corpora de Estudo:

- *Dubliners* (1914), de James Joyce
- Um dos quinze contos

■ Momentos da avaliação da tradução literária:

- Análise do texto original
- Estudo comparado do original e das traduções



A Definição de um Corpus de Referência (Berber Sardinha, 2004: 97-101)

- Também conhecido como corpus de controle, funciona como termo de comparação para a análise.
- Sua função é fornecer uma norma com a qual se fará a comparação das frequências do corpus de estudo.
- As palavras cujas frequências no corpus de estudo forem significativamente maiores segundo o resultado da prova estatística são consideradas chave.



Corpus de Referência: Características Gerais

- O corpus de referência não deve conter o corpus de estudo;
- Um corpus de um gênero específico, igual ao daquele do corpus de estudo, tende a filtrar os elementos genéricos (relativos a um mesmo gênero) comuns;
- Um corpus de *referência geral*, incluindo vários gêneros, não excluirá as palavras genéricas (relativas a um gênero);
- Tamanho mínimo necessário para resultados significativos (mínimo de 2 a 5 vezes o corpus de estudo)



1º Momento: Análise do texto original

A Construção de um Corpus de Referência Literário: *Crítérios:*

- Escolha de um corpus de referência de características genéricas (gênero **conto**) – Compilação de contos
- Escolha de autores contemporâneos e de mesma estatura literária: (James Joyce :1882-1941)
 - { Virginia Woolf (1882-1941)
 - { Katherine Mansfield (1888-1923)
 - { D. H. Lawrence (1885-1930)



A Construção de um Corpus de Referência Literário:

Composição do corpus de referência:

Autores	nº de contos	Total de palavras
Virginia Woolf	23	53.659
Katherine Mansfield	12	79.087
D. H. Lawrence	17	79.845
TOTAL	52	212.591

Obs.: Corpus de Estudo: 67.936 palavras
(3,13 vezes menor)



O Programa *WordSmith Tools*

Mike Scott, versão3, 1996

A ferramenta *KeyWords*

- Compara duas listas de palavras e fornece uma lista de palavras-chave, cujas frequências são estatisticamente diferentes no corpus de estudo e no corpus de referência;
- Chavicidade: Medida de quão “chave” é a palavra considerada;
- Palavras-chave positivas/**negativas**: ocorrem mais/**menos** frequentemente do que o esperado aleatoriamente numa comparação com o corpus de referência.

Exemplos da Lista de palavras-chave:

Corpus de Estudo: *Dubliners*

Corpus de Referência: Woolf, Mansfield, Lawrence

5 primeiras palavras (positivas):[p=0,000001]

N	WORD	FREQ. DUBJJ.LST	%	FREQ. REFCOR.LST	%	KEYNESS
1	MR	573	0,84	173	0,08	916,1
2	GABRIEL	141	0,21	0		400,1
3	AUNT	125	0,18	31	0,01	216,3
4	HIS	1.158	1,70	2.158	1,02	193,0
5	KERNAN	66	0,10	0		187,2

Exemplos da Lista de palavras-chave:

5 últimas palavras (negativas)

N	WORD	FREQ.	DUBJJ.LST %	FREQ.	REFCOR.LST %	KEYNESS
129	MOTHER	30	0,04	295	0,14	48,6
130	IT	581	0,86	2.808	1,32	100,9
131	HER	790	1,16	3.674	1,73	112,5
132	OH	3		320	0,15	152,0
133	SHE	695	1,02	4.471	2,10	376,6



Exemplos da Lista de palavras-chave:

O que as palavras-chave positivas revelam:

Definição de áreas temáticas:

- Música (tenor 45,4; concert 43,8)
- Alcoolismo (bottles 34,9; whisky 28,8)

O que as palavras-chave negativas revelam:

Ausência de grupos semânticos:

- Emoção (dear -35,9; cried -24,2)
- Cores (white -31,5; blue -27,9)
- Natureza (garden -36,7; sea -24,2)



2º Momento:

Estudo comparado do original e das traduções

- **Novos corpora de estudo:** qualquer dos contos de *Dubliners* e respectivas traduções
- **Novos corpora de referência:** Os outros 14 contos de *Dubliners* (*dub-conto*) e os de cada tradutor
- **Expectativas:**
 - Características particulares do conto
 - Características particulares dos tradutores

Exemplos de Comparação

Palavras-chave

Conto: “A Little Cloud” (**alitcl**)

Tradutores: Hamilton Trevisan (**upn-ht**)

José Roberto O’Shea (**upn-os**)

Palavras-chave (**alitcl x dub-alc**):

N	WORD	FREQ.	ALITCL.TXT %	FREQ.	DUB-ALC.LST %	KEYNESS
1	GALLAHER	48	0,95	0		251,5
2	CHANDLER	41	0,81	0		214,8
3	IGNATIUS	38	0,75	0		199,0
4	LITTLE	57	1,13	153	0,24	76,2
5	LONDON	12	0,24	4		45,4

Exemplos de Comparação

Palavras-chave

Comparação upn-os x d-upn-os:

N	WORD	FREQ.	UPN-OS.TXT %	FREQ.	D-UPN-OS.LST %	KEYNESS
1	GALLAHER	58	1,18	0		300,6
2	CHANDLER	44	0,90	0		227,9
3	LITTLE	40	0,82	0		207,2
4	IGNATIUS	38	0,78	0		196,8
5	AMIGO	23	0,47	23	0,04	58,9

Comparação upn-ht x d-upn-ht:

N	WORD	FREQ.	UPN-HT.TXT %	FREQ.	D-UPN-HT.LST %	KEYNESS
1	GALLAHER	43	1,06	0		228,4
2	CHANDLER	36	0,88	0		191,1
3	LITTLE	33	0,81	0		175,2
4	LONDRES	11	0,27	3		44,2
5	ANNIE	8	0,20	0		42,4



Exemplos de Comparação

Palavras-chave

Comparação de **upn-os** com **upn-ht**:

(Utilizando o comando “compare two lists”, da ferramenta Wordlist)

N	WORD	FREQ.	UPN-OS.TXT %	FREQ.	UPN-HT.TXT %	KEYNESS
1	IGNATIUS	38	0,78	5	0,12	23,1
2	PRA	13	0,27	0		15,7
3	TÁ	7	0,14	0		8,5
4	TE	6	0,12	0		7,3
5	VOLTA	6	0,12	0		7,3



Conclusões

Para a análise do texto literário:

-  Determinação de áreas temáticas
-  Determinação de “ausências” no texto
-  Determinação das particularidades de um sub-texto/corpus

Para a análise/avaliação de traduções literárias:

-  Verificação da conservação das palavras-chave do original
-  Determinação de características de diferentes tradutores

This document was created with Win2PDF available at <http://www.daneprairie.com>.
The unregistered version of Win2PDF is for evaluation or non-commercial use only.